A Progressive Universal Noiseless Coder

Michelle Effros, Philip A. Chou, Eve A. Riskin, and Robert M. Gray

Information Systems Laboratory, Stanford University, Stanford, CA 94305-4055, email: effros@isl.stanford.edu
Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304, email: chou@parc.xerox.com
Dept. of Electrical Engineering, FT-10, University of Washington, Seattle, WA 98195, email: riskin@ee.washington.edu
Information Systems Laboratory, Stanford University, Stanford, CA 94305-4055, email: gray@isl.stanford.edu

We describe an adaptation of Itoh and Kawabata's universal noiseless coder that allows for progressive transmission of images. The system is based on a tree structure, and codewords stored at internal nodes of the tree allow for early reproductions of the input image. When the encoder reaches a leaf of the tree, it continues transmitting until the compression is lossless. Compression results compare favorably to Ziv-Lempel coding of both images and finely quantized Gaussian sources.

Motivated by the need in the medical image community to compress images robustly, transmit them progressively, and ultimately reconstruct them noiselessly, we have adapted Itoh and Kawabata's universal noiseless coder[1] to the task of progressive transmission. In our system, the decoder reconstructs increasingly better reproductions of the transmitted image as bits arrive. This allows for early recognition and is useful in telebrowsing of image databases. Most work in progressive transmission to date[2] has involved the transmission of images compressed with a lossy technique or with no compression at all. Our system compresses images losslessly at rates close to their entropy rate regardless of their source (e.g., imaging modality).

Itoh and Kawabata's universal noiseless coder is based on the principle of minimum description length. First the encoder describes to the decoder a model of the distribution of the data; then it describes the data using the model. Of all models in a given class of models, the model chosen is one that minimizes the total description length. Itoh and Kawabata use as their class of models variable-length binary tree structures with probabilities stored at each leaf. The trees recursively split the input space in half until the distribution is roughly uniform across each bin. The leaf probabilities allow the index of a bin to be entropy coded. The exact element within a bin (typically a point in a k-dimensional rectangular lattice) is transmitted using a fixed rate code matched to the bin volume. Itoh and Kawabata have shown that when the trees are optimally pruned to minimize the total description length, the total description length per letter approaches the source entropy under mild smoothness conditions on the source density.

Our compression system is similar, but instead of storing at each leaf the probability of reaching that leaf, we store at each internal node the probability of going to a left or right child from that node. Storage of these intermediate node probabilities is equivalent to Itoh and Kawabata's storage of leaf probabilities, but allows for progressive transmission when an arithmetic code is used to transmit the binary path maps for the data. The node probabilities are economically encoded using a method of Langdon and Rissanen[3] and have at most 3.5% redundancy. Also stored at each node is the centroid of the data in the node. The centroid is encoded using the same fixed rate code that is used to noiselessly describe the data in a leaf. Because of this additional stored information, our tree description length is larger by a constant factor than the tree description length of Itoh and Kawabata. Nevertheless, the description of our tree still contributes a vanishingly small fraction of the total description length as the data length increases. Hence, as a noiseless coder, our system inherits the universal properties of Itoh and Kawabata's coder.

To use our tree structure for progressive transmission, we prune back our "noiseless" tree using the pruning algorithm of [4, 5] to obtain a sequence of nested subtrees at a decreasing sequence of rates (and an increasing sequence of distortions). Each tree in the sequence is optimal in the sense that it codes the data to the lowest possible distortion among all the subtrees of the original "noiseless" tree having the same or lower total description length. This sequence of trees defines the order in which the bits of the path maps describing the data should be sent for progressive transmission. More precisely, if $T_0 \prec T_1 \prec T_2 \prec \cdots$ is the sequence of nested trees in order of increasing rate, then the overall encoding at stage $i+1$ is the overall encoding at stage i, followed by an encoding of tree T_{i+1} given tree T_i, followed by an encoding of those data whose path maps extend beyond T_i to T_{i+1}, using an arithmetic code to encode the path map extensions. The encoding of tree T_{i+1} given tree T_i may be accomplished by transmitting the sequence of nodes that when split produce T_{i+1} from T_i. Each node in this sequence can be indicated by its path map in the current tree. The overhead for transmitting these path maps becomes negligible as the data length increases. When the decoder receives the path map extension for a datum, it reconstructs the datum as the centroid of the appropriate leaf. However, if the path map terminates in a "noiseless" leaf, the decoder reconstructs the datum noiselessly from a fixed rate code (matched to the leaf), which immediately follows the path map. Thus progressive transmission is extended naturally from lossy to lossless compression.

Results on Gaussian sources that are uniformly scalar quantized to 8 bits show that the per letter total description length goes to the first-order entropy as the number of data points N increases. As the block size k increases, the per letter total description length goes to the entropy rate. We get similar results for medical image sources and images from the USC database. In all cases, our noiseless compression rates are comparable to Ziv-Lempel, but we have the advantage of progressive transmission.

References

[1] S. Itoh and T. Kawabata. A universal data compression algorithm with parameter variance estimates. In *Proceedings of the VII-th Symposium on Information Theory and Its Applications*, pages 239–244, Nara, Japan, December 1985. In Japanese.

[2] K-H Tzou. Progressive image transmission: a review and comparison of techniques. *Optical Engineering*, 26(7):581–589, July 1987.

[3] G. G. Langdon and J. Rissanen. Compression of black-white images with arithmetic coding. *IEEE Transactions on Communications*, 29(6):858–867, July 1981.

[4] P. A. Chou, T. Lookabaugh, and R. M. Gray. Optimal pruning with applications to tree structured source coding and modeling. *IEEE Transactions on Information Theory*, 35(2):299–315, March 1989.

[5] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. The Wadsworth Statistics/Probability Series. Wadsworth, Belmont, California, 1984.